Computing News

News from the Computing Division Fermi National Accelerator Laboratory

VOLUME XXII, NUMBER 2, AUGUST, 1994

CRADA

Fermilab has entered into a CRADA (Cooperative Research and Development Agreement) between the Computing Division and a major computer company. CRADAs contribute to the goal of transferring technology developed at Fermilab into the commercial sector, with benefits to both parties and to the nation's technological strength.

An agreement executed with CRAY Research, Inc. aims at porting the Canopy software framework to their CRAY T3D massively parallel system. The effort involves members of the HPPC department and uses the T3D system at the Pittsburgh Supercomputing Center. This technology transfer will open the Canopy tools for use in many other scientific and engineering applications.

Canopy was developed at Fermilab by the Computing Research and Development Department and the Research Division Theoretical Physics Group and has been in use in the ACP-MAPS system since 1991. Canopy provides a natural environment for solving grid-oriented problems. It includes tools such as links, sets and fields over grids, and subroutines for doing numerical analysis operations. The goal of Canopy is to allow scientists to use massively parallel systems for a broad class of applications without having to become expert in any particular system or with parallel-programming techniques. Canopy uses the natural parallelism in most grid-oriented problems automatically, isolating the user from the multi-node nature of the machine. By using the grid-oriented concept of performing the same operation on many sites of the grid simultaneously, it encourages structured programming as well as efficient use of the parallel-processing machine.

Mark Fischler, x4339, mf@fnal.gov

CDF and D0 Computing Upgrades

The ongoing collider run promises a large increase in the amount of data that CDF and D0 will collect and analyze. Enhancements to the computing of CDF and D0 are being made to cope with the increased data. The major changes which are immediately needed are increases in UNIX farm reconstruction capability and to the central splitting, file-serving, and analysis machines.

What's Inside

The CDF farms have been augmented with the addition of a SGI Challenge XL (4 R4400 150 MHz processors) with approximately 100 GB of disk and 12 double-density exabyte tape drives. The Challenge is serving as an I/O node for up to 64 worker nodes (SGI 4D/35) in 4 ethernet segments and as a splitting system for the output streams coming from the farm. This system replaces a SGI 4D/420 and a set of VAXstation 4000/90's that were used as an I/O node and a splitting system for RUN1A. In addition, an IBM RS6000/580, RS6000/530H and RS6000/590 with approximately 200 GB of disk have been or will be soon added to the CDF farms. These nodes serve as I/O nodes for up to 48 worker nodes and as a splitting system for the output streams coming from the farm. The system is not quite at its full size but is being expanded as needed to handle the increasing luminosity from the Tevatron.

The D0 farms have been increased in size by replicating the basic unit of one SGI 4D/420 I/O node with 12 GB of disk and 7 double-density exabyte tapedrives. Each I/O node has 24 worker nodes (typically R3000 Indigos) attached in 3 ethernet segments. The farms have been increased from 24 to 74 worker nodes (the third system has 26 worker nodes) and can be further expanded by adding more I/O nodes. D0 has 2 standalone splitting systems, each one consisting of a SGI Crimson with 6 tapedrives and 12 GB of disk. These are used to split the STA (full size) output of the farms into "streams" of physics samples. The DST's written on the farms are split on the D0 file system (D0FS).

The CDF central analysis systems are both being upgraded in preparation for the RUN1B data. The central UNIX system (cdfsga) is to be upgraded from a SGI 4D/480 (rated at approximately 240 MIPS) to a SGI Challenge XL with 8 150 MHz R4400 processors (rated at approximately 800 MIPS), giving a sizable increase in the CPU capability as well as a clear path for expansion. The Challenge XL can be upgraded to 36 processors and those processors can be upgraded to faster R4400's or the new R8000 chip. 170 GB of disk will be added to cdfsga as part of the upgrade. The plot of CPU usage on the CDF VMS and UNIX analysis systems (see Computer Usage section) shows that the UNIX system, after a slow start, has become an important part of CDF analysis computing.

The central VMS system (FNALD) is to be enhanced by the addition of two DEC AXP (Alpha) 2100 A500 (Sable). Each Sable will contain 2 Alpha CPU's and will add approximately 440 MIPS to the current cluster. 250 GB of disk will be added as part of this acquisition. In addition, the VAX 6550 (FNALO) which is to be decommissioned from FNALV will be moved and attached to FNALD.

The D0 analysis systems are to be augmented with a new SGI Challenge L system and with many enhancements to the central and widely distributed D0 clusters. The Challenge L is being ordered with four 150 MHz R4400 processors and with 170 GB of disk. The cluster enhancements include networking upgrades to allow faster and more reliable data transfers, memory upgrades for various systems, additional disk space, DLT and exabyte tapedrives and software enhancements for the D0 systems. All of these enhancements are meant to increase the performance of current D0 systems for RUN1B data.

Steve Wolbers, x3950, wolbers@fnal.gov

C D

Transitioning Oracle from FNALV to FNALU

The Oracle license on the FNALV cluster will be moved from FNALV to FNALU in late fall. This move is in accordance with the policy to transition applications from VMS to vendor neutral cost effective Unix platforms. In addition, Oracle performs better when running on IBM, SUN, or SGI servers and Unix platforms provide better environments for Oracle's GUI line of products such as Oracle Forms 4.0, Oracle Reportwriter 2.0, and Oracle*Case. All Oracle products, currently offered on the FNALV cluster will be available on the Unix cluster. These products include:

RDBMS version 7.0
SQL*Plus
SQL*Forms 3.0
SQL*Report 1.1
Pro*Fortran
Pro*C
SQL*Net for connectivity to other platforms

In addition, we will add the up-to-date GUI products, Oracle Forms 4.0 and Oracle Reportwriter 2.0.

A conversion plan that will provide for a smooth transition and impose minimal inconvenience on our users has been designed. During this transition period, Oracle will continue to run on the FNALV cluster. Users will have the opportunity to convert to Unix and the Computing Division will provide assistance. We expect that the complete transition to occur in three months.

Jack Stein, x8273, stein@fnal.gov

CAP Project

The CAP project is moving toward its short-term goal of providing high-throughput access to large datasets, for the "data-mining" phase of experiment analysis. The purpose of CAP is to facilitate scanning through hundreds of Gbytes of summarized data to select a small fraction for further analysis by the user physicist. It will also provide smooth access to the full data associated with events — many Tbytes of data, residing in a robotic automated tape library.

An early prototype of the data-mining and extraction system will be available in early August. Implemented on the IBM 9076 (SP1 and SP2) computer, it will support "queries" through more than 100 Gbytes of on-line summary data, and access to about a Tbyte of robotic tape storage for full data of a subset of those events. Early users will include selected D0 physicists, who will have rapid access to DST data from run 1a, and possibly CDF physicists who may access PAD and/or DST data from accumulating "exotics" streams and designated "hot events" collections.

The sort of queries supported ranges from simple selections based on attributes of a single physics object (the kind of n-tuple oriented selection supported by **PAW** and **PIAF**), through more complicated selections based on attributes of multiple objects in an event. For example, a single pass can select events based on properties of an electron and a jet, and/or based on the presence of two muons with particular properties. The following is a sample query:

```
Muon#1.E > 30. && Muon#1.E < 40. && Muon#2.Pt >
10. && Electron.E > 15.
```

The user can specify whether to extract the full event data for selected events (typically on tape), just the summary data kept on disk, or other options. The optimization emphasis is on rapid scanning of the selection data, and extracting the full or summarized events selected, rather than on forming histograms of properties of the selected events on the fly (although the latter, optimized in CERN's new **PIAF**, is also supported). This very early prototype should yield information as to what functionality is important to the users, and how usage patterns change when physicists are presented with the capability to rapidly pass through their large event sets. It should also provide a "proof of principle" that this kind of architecture is suitable for useful HEP data analysis.

The prototype query engine is based on an object-oriented library supporting the concept of *persistent pointers*. These are pointers to objects (events, or components of events such as electrons, muons, and jets) which persist across jobs: They reside on disks and/or in tertiary tape storage. **Ptool**, invented by B. Grossman at UIC and employed in the PASS project, provides a light-weight set of tools implementing persistent pointers in C++.

The query strategy is to divide the data into *stores* each representing one type of physics object. For example, there might be an electron store and a photon store. Each store is kept on disk in a manner facilitating rapidly sweeping through that data. Since a typical query is concerned with only a few physical objects to select events, this strategy implies that only a small fraction of the summarized data must be brought in off disk.

Experimental data exists in varying native formats (ZEBRA or YBOS banks for D0 and CDF). The query tool needs some fraction of the data — the parts to be organized for rapid scanning — to be translated into C++ classes representing the objects to be scanned (electrons, photons, and so forth). We call this the *cardinal data*: It will reside on disk for rapid scanning. Creation of headers defining these *physics objects* for use on CAP, and of an unpacking routine to extract the objects from the native databanks, has been done in the case of D0. D0 physicists have also implemented a re-packer, to reconstitute desired parts of extracted events in the native ZEBRA format.

(CDF has independently come up with a similar strategy for data mining: A "multifile format" for PAD data -- e.g., an electron list, a muon list, etc. Given the identification of those files, it is straightforward to create the classes defining physics objects, and to go from those files to the needed physics objects.)

The goal is for scans through data sets of a couple of hundred Gbytes to routinely proceed in less than an hour. Summary data kept on disk and full data for a fair number of "hot" events kept in a disk cache should be available quickly. Extraction of a substantial number of selected complete events (assuming the full reconstructed data must be read in off tape) would proceed overnight, to be able to group requests to minimize total tape activity. (On current systems, such full-scan data mining sweeps can take weeks.) The first prototype, although not fully optimized, should provide some idea of how close the system is to meeting these goals.

The other immediate thrust of CAP is smooth access to many Tbytes of data in an automated tape library. A large robotic system, controlled by Unitree hierarchical storage manager software, is currently connected to the IBM system. In the fall, new technology tape drives will be installed in a similar robot, providing access at low enough per-Tbyte costs to obtain the tens of Tbytes needed to keep a substantial body of full event data conveniently accessible. This will of course be integrated with the event selection/extraction mechanism for data mining, and will also provide general access to specific events or runs of data to the lab community.

Around the end of the year, CAP should have a prototype production system, with reasonably optimized data mining capabilities, a broad variety of features needed for such analysis, and access to a trial large robotic system. This can provide an important new facility for post-event-reconstruction data analysis.

This project involves efforts by all groups within the HPPC Department. In particular, significant contributions are being made by Kirill Denisenko in utilizing the object-oriented techniques.

Mark Fischler, x4339, mf@fnal.gov

White Pages Project

Fermilab has joined the ESnet/NERSC White Pages Project which provides white pages directory services to members of the Energy Science Network. The white pages project was created to allow members of the Energy Science Network the ability to locate information about a colleague over the Internet, similar to a large online phone book. Anyone with an account on the lab mailserver (FNAL) has their lab extension, Internet mail address and US mail address made available to the Internet via the project.

There are a number of ways to retrieve the information and we will describe them all in the next issue of Computing News. One way is via WWW using the url: `http://www.es.net/` Once there, select the **White Pages** icon under the **Other ESnet Information Services**.

Jack Schmidt, x4060, schmidt@fnal.gov

UNIX

UNIX Application Support Group Product News

As part of the recent Computing Division re-organization, the UNIX Application Support Group (UAS) was formed as a portion of the Operating System Support (OSS) Department. As the group's name implies, its focus is in providing support for a variety of UNIX Application Software. This includes locally developed software such as **ups**, **rbio**, and **juke**, free software on the Internet such as **emacs** and **ghostview** and commercial products such as the **edt** and **tpu** editors. The group doesn't provide any Physics applications - these are still provided by the Physics Analysis Tools (PAT) Department. In addition, there are many other providers of UNIX products at Fermilab, including but certainly not limited to the Online Support Department and the High Performance and Parallel Computing Department of the Computing Division.

The status of products supported by UAS is posted on a regular basis to the newsgroup fnal.comp.unix. In addition, this information is available via WWW at URL: `http://www-oss.fnal.gov:8000/uas/release.txt`.

The purpose of this article is to highlight some of the recent activities of the group.

Support of Other Command Shells

In July, UAS began providing official support for two additional login shells: **tcsh** and **bash** as part of v1_0 of the **shells** products. **tcsh** is a superset of **csh** and contains additional interactive features that are preferred by many users. **bash** is a ksh-like shell that is nearly, if not completely, POSIX-compliant. Both of these shells have been requested by a variety of users. The products is available for AIX, IRIX, and SunOS.

Both of these shells should work as part of the standard Fermi UNIX environment as do Bourne shell (**sh**), **csh** and **ksh** provided by the vendors. Users should use whatever shell they are most comfortable with for interactive use. The Bourne shell is often preferred for scripts as this is universally available and is a more consistent syntax for complex script writing.

In September, a revised version of the product was released that also includes **perl** and the GNU version of **awk** (**gawk**). Although these are not interactive shells, they are both powerful scripting languages and thus fit the flavor of the **shells** product.

New version of TeX

A new version of TeX, v3_1415 has been released. The main feature of this release is an updated version of macros that are already being used at some sites such as CERN. In addition, **dvi2tty** (converts dvi output to ASCII) and the **makeindex** product are also being included.

With this version of TeX, a packaging change is being introduced. The TeX product is being split into two products, **tex** and **tex_files**. Users will interact with the **tex** product which contains all of the binaries. The **tex** product will rely on the **tex_files** product which contains platform-independent files such as fonts. The **tex_files** products only needs to be installed once (even on a UNIX cluster containing machines of different architectures) thus saving nearly 40 Mbytes of space per architecture type.

The local TeX documentation (PU0066) is also being updated and will reflect both UNIX and VMS usages.

New release of flint

Version 3.0 of **flint** (Fortran-lint) is now the current version and is available for all systems that previously had access to Version 2.90. Release notes are available as part of the **flint** package.

Additionally, the serving of the **flint** product has been moved to FNALU (specifically fsgm01) from cdsgi. System managers should take note of this.

Upon declaring Version 3.0 current, system managers should make a clean break from Version 2.90. Version 2.90 will continue to be available and supported through March 1995. After that time, Version 3.0 will be the only supported and available version of **flint**.

If you need instructions on how to migrate from version 2.90 to version 3.0, please send mail to `uas-group@dcdsv0`.

Updated ghostview/ghostscript support

Currently, **ghostview** and **ghostscript** have been released each as separate products. Beginning with **ghostview** Version 1_5, the two products are combined and released as one product, **ghostview**. By installing **ghostview**, users will be able to use either **ghostview** or **ghostscript**, or both.

UAS will update its **ghostview** product when either new versions of **ghostview** or **ghostscrip**t become available. The

version of the combined **ghostview** will correspond with the version of the individual **ghostview** product. For example, if the latest version of the individual product **ghostview** is 1_5, the combined version will be **ghostview** 1_5. A new release of the individual product **ghostscript**, will result in a "sub" release of the combined **ghostview** product, that is, **ghostview** v1_5a.

Support of DLT-2700 Tape Libraries

juke is a product that provides command line and C library calls to control the basic functions of a tape jukebox. Originally the package was written to support the Exabyte 10 and 120 tape robots. With the release of v3_0 of the product, the DLT-2700 tape library is also supported.

Support of GNU Software Development Utilities

For quite some time, a variety of customers have requested support for the GNU compilers. UAS has plans not only to support the GNU compilers, but a variety of other GNU software development utilities. The support plans call for the C and C++ compilers, gnu make (**gmake)**, gnu yacc and lex (**bison** and **flex)** gnu awk (**gawk**), and the gnu **debugger**. These GNU products will be packaged into several UPS products so that customers may obtain a subset of the entire GNU software development environment. The GNU products consist of three separate products: **gcc**, **libgpp**, and **gtools**. All three of these products are available as test.

Matt Wicks, x8083, wicks@fnal.gov

Etta Burns, x8300, ettab@fnal.gov

C D

New Way to Get UNIX Products

Traditionally, the supported method for obtaining UNIX products from the Computing Division has been through the use of **upd** (UNIX Product Distribution). The Computing Division has maintained a repository of **ups** products accessible through **upd** on the machine fnsg01.

upd is connected with a **ups** database that allows you to obtain a list of products available, determine version numbers, determine which version is current, etc. It also allows you to retrieve the associated tar file for the product, unwind the file, and declare it to your local **ups** database. When using the **`ups declare`** feature of **upd**, product dependency information is provided as well. All of these functions are accomplished via a menu interface. Additional features of **upd** can be found in PN435, *UNIX Product Distribution User's Guide* available in the Computing Division Library.

Although **upd** has quite a few advantages, it does have some disadvantages - most notably it requires obtaining, installing and configuring the **upd** product before you can obtain any software. In addition, **upd** requires `.rhosts` registration for all systems that wish to obtain software.

In order to address the problems with **upd**, the Computing Division started the EZ-UPD project. The goal of EZ-UPD was to provide enhancements to UNIX product distribution, without eliminating the advantages of the current **upd** system. The end result of this project has been to create an anonymous ftp repository for all Computing Division UNIX software.

Anonymous ftp is a de-facto standard mechanism for ob–taining software on the Internet. All UNIX systems supported at the lab already come with the ftp software, thus access to the software is automatic. Access to the software via **upd** is still supported. In fact, the **upd** repository and the anonymous ftp repository point to the same location.

In order to gain access to the software, you simply execute the command: `ftp ftp.fnal.gov`. Specify a user name of either `anonymous` or `ftp` and provide your e-mail address as the password. The software packages are organized by operating system type (`GENERIC_UNIX`, `AIX`, `IRIX`, `SunOS`, `ULTRIX`, `HP-UX`, and `OSF1`).

Registration is still required to access the software via anonymous ftp. However, all nodes in the `.fnal.gov` domain are automatically registered. Off-site nodes must fill out the same registration form required for **upd** access. Registration for **upd** results in registration for anonymous ftp and vice-versa. The registration form is available in the `FORMS` folder of **INFO** on VAX, the `fnal.announce.forms` newsgroup, and anonymous ftp on `ftp.fnal.gov`.

The `ftp.fnal.gov` machine is running an enhanced version of the ftp server. This server presents a greeting message after initial login which provides basic instructions. The greeting message refers to the readme file which provides additional instructions as well as the index file which provides a complete list of available files.

Another feature of the anonymous ftp server is on-the-fly compression. The tar files are stored uncompressed on the server for compatibility with **upd**. However, you may retrieve a compressed version of the tar file in order to conserve network bandwidth and local disk space. In order to request a compressed tar file, append either a `.Z` (for standard compressed format) or `.gz` (for gzip format) suffix at the end of the file name. For example, instead of executing the ftp command

```
get upsv3_2aIRIX.tar
```

you would execute

```
get upsv3_2aIRIX.tar.Z    or
get upsv3_2aIRIX.tar.gz
```

In addition to using the standard ftp interface, the software may be obtained using a WWW browser such as **xmosaic** by specifying as your URL: `ftp://ftp.fnal.gov/`. WWW provides a point and click interface through the directory structure. Clicking on a tar file results in the file being transferred to your local machine. The greeting message and on-the-fly compression features of the ftp server are not accessible when using WWW.

Matt Wicks, x8083, wicks@fnal.gov

FNALU Upgrades

We plan to increase the capabilities of the central UNIX system FNALU this year by a combination of hardware and software additions. First, two major hardware increases will be made. A Sun SPARCserver 20 Model 514MP will be added to the system. This is a 4-processor 50 Mhz Sun system. The SUN will give FNALU a third flavor of UNIX (IBM's AIX and SGI's IRIX are the other two). Second, a SGI Challenge XL with 4 150 MHz R4400 CPU's will be added to FNALU. The Challenge allows increased interactive and batch capabilities on FNALU and provides a substantial upgrade capability. Additional CPU memory, disk and tape can all be added to this system as the needs arise.

Once the new hardware is in place, FNALU will consist of 3 SGI computers (a 4D/420 a 4D/480 and the Challenge XL), 2 IBM RS6000's (models 560) and the SUN Sparc 20. The filesystem for FNALU, though not properly part of the system itself, will also be upgraded as part of a plan to expand the use of AFS at Fermilab. This is described in a separate article.

Various software additions to FNALU are being made to improve the capability and usefulness of the systems. The Oracle license and databases that currently reside on FNALA will be moved to FNALU. Standard compilers (Fortran, C and C++) are being purchased for many of the FNALU systems. The SGI CASEVision and IDO bundles are being purchased for the Challenge. The Sun system will have full debugging and code management tools, along with a parallelizing Fortran compiler. In addition, software that allows PC windows applications on the Challenge are being purchased. These initially include Microsoft Excel, Microsoft Word and Microsoft Power Point. Other windows applications can be purchased and licensed to run on this system if the need arises. Some Windows applications will also be available on the Sun system via the WABI windows emulator. In addition, several licenses of Framemaker are being purchased. These will be accessible from the SUN, SGI and IBM portions of FNALU.

These hardware and software upgrades to FNALU will help to make the cluster more useful to the Fermilab community, and will increase the UNIX computing power available to experimenters.

Steve Hanson, x8043, hanson@fnal.gov

AFS Expansion

Several plans for expansion of the use of AFS at Fermilab have been recently announced by the Computing Division. This includes the expansion of AFS file serving capabilities on the FNALU cluster, and the installation of AFS client software at various locations at Fermilab to be able to make use of central file serving resources.

As part of the current FNALU expansion, the three IBM AFS file servers will be replaced with four Sun SPARCserver 5 systems. This will allow the current IBM file servers to be reused as local AFS file servers at locations within the lab that would benefit from local file serving. This will reduce the load on the central file servers and lower the network traffic required to serve software products and home directories throughout the laboratory. Additionally, 60 Gigabytes of additional disk space will be added to the AFS file servers during the expansion.

Another part of the AFS expansion plan is to provide all of the Computing Division supported UNIX products in AFS space. By using the multiple database feature of **ups**, users can have access to both the products in AFS space as well as their local products. The products in AFS space will be replicated across several AFS servers. Thus, access to these products will be improved and not affected by the failure of an individual server.

AFS client software will also be installed in selected locations throughout the lab as we expand the use of AFS beyond the FNALU cluster to distributed work clusters. These systems have been chosen as having particularly good uses for the AFS file system, and as being representative of some of the problems that will be faced as AFS is deployed. These systems will then begin to make use of the AFS file system facilities to obtain central software products, and may also use AFS as user file storage.

After the initial pilot installations, it is our intention to make AFS client software available through an automated installation procedure for all of the supported UNIX systems. This software will be installed according to user demand. Eventually, any UNIX system at Fermilab will have access to AFS files both here and at other locations on the Internet.

Steve Hanson, x8043, hanson@fnal.gov

Communications

Fermilab ISDN Pilot Project

ISDN (the Integrated Services Digital Network) is changing the way Fermilab scientists work at home. ISDN, a technology developed by the world's phone companies over the last 10 years, uses a normal pair of telephone wires to deliver a digital connection. With basic rate ISDN (BRI), a pair of wires carries up to 128 Kbps of user information. For work-at-home application, this marks a dramatic increase in speed.

Since most computers at home (PC, Mac, X terminal) can only handle 20-30 Kbps through their serial ports, ISDN equipment vendors have focused on demand bridging. With demand bridging, a Fermilab Ethernet LAN is extended to the home. The user at their home machine can mount files, bring up multiple X windows, and share printers exactly as if they were at work. Also, the ISDN bridge connects to the home machine via Ethernet, thus bypassing the slow serial interface. Finally, the

exact same configuration supports TCP/IP, Appletalk, and DECnet access.

The HEP Network Resource Center at Fermilab has been experimenting with ISDN for several years. Until recently both ISDN service and equipment was expensive and rare. Within the past year, however, Ameritech has made a commitment to deliver BRI anywhere in the Chicago area and equipment prices have fallen dramatically. BRI service is available to virtually all Fermilab employees homes for $40-$60 per month plus ISDN equipment costs of about $600.

Based on the availability and cost of ISDN service and equipment, the Computing Division began a pilot project in late 1993. Fourteen users were chosen from throughout Fermilab. By June 1, 1994, all 14 lines were installed and working. The fourteen users became more since several households had more than one Fermilab employee. Each household was supplied with a Combinet CB-400 bridge and a Northern Telecom NT-1 to connect to either their computer or X terminal.

The remote equipment is configured so that Fermilab is automatically dialed as soon as traffic is present on the home Ethernet. For example, if a user boots a Macintosh or brings up a telnet window on a X terminal, the connection will be established a few seconds, without user intervention.

Security is provided by passwords and callback as is the case now with modem access. The difference between this and "normal" telephone/modem connections is that all of this is transparent to the user and happens within a matter of seconds.

Initial response from the testers has been positive. In particular, the testers like the fast call setup time and the high bandwidth with its resultant "snappy" response time even for graphics.

The Computing Division is working on the architecture for the expansion of the project. Plans are being prepared to extend this project to a general service similar to that now provided over standard telephone lines with modems.

Bill Lidinsky, x8067, lidinsky@fnal.gov

C D

ISDN Access Coming To The Village

The Village dorms have long been short on communications ability. A few ASCII serial terminals are connected through multiplexors to the port selectors, using dedicated analog telephone lines. This is far less than adequate for anyone trying to do useful work without taking a hack into their office or lab. Thanks to the new high-speed analog and digital modems, many people now have better communications from their home than is available to the dorm residents. However, the high cost of trenching new communications lines in the Village has prevented any significant upgrade of the dorms.

With the proven success of the ISDN pilot project in providing high-speed (128 kb/s) digital access to Fermilab (see accompanying article), a low-cost alternative is now available to significantly improve the situation. The Computing Division is planning to provide and install high-speed digital ISDN modems, identical to those used in the ISDN pilot project, for the dorms. In order to make efficient use of the higher bandwidth, the Computing Division is also reviewing the computing equipment (terminals, etc.) available in the dorms and will make recommendations on upgrades to these as well.

Mark Kaletka, x2965, kaletka@fnal.gov

C D

Upgrades To Dial-in Services

New Faster Modems Supported

In mid to late September, all modems in the direct-dial and call-back modem pools will be upgraded to support V.FAST (V.FC) connections. This new protocol will support basic carrier rates up 28,800 bits per second and, when combined with data compression, data rates of two to four times the basic rate. (Actual observed maximum rates can vary and are highly dependent on the quality of the phone lines in use.) V.FAST is an intermediate step in providing all the features of the proposed V.34 international standard which is now under consideration. The Fermilab modems will be upgraded to full V.34 compliance as soon as the new hardware becomes available in early 1995. The V.FAST upgrade will support all the existing V.32 modems now in use which provide service up to 14,400 bits per second. However, some care is required when purchasing new desktop or add-in modems for home use if they are to be compatible with the faster data rates possible with the intermediate V.FAST upgrade.

In order to realize the full benefits of our V.FAST upgrade look for a modem which is V.FAST compliant or which uses the Rockwell V.FC chip set. Manufacturers include Microcom, Hayes, Practical Peripherals, Supra, Zoom and others. Modems from AT&T and Motorola use different chip sets and are NOT compatible with our V.FAST upgrade. All V.FAST compliant modems are supposed to be compatible with the future V.34. Most manufacturers promise an upgrade path to the V.34 standard for a fee. The fees vary so ask ahead of time. Expect to pay $200 to $300 for a V.FAST modem although some ISA add-in cards have been seen advertised for as little as $130. If prices continue to fall it could render the price of an upgrade to V.34 moot.

Appletalk Remote Access Available

In early August, Apple Remote Access Protocol (ARAP) V1.0 was made available on the modem pool terminal servers. This is in addition to the other connection protocols already supported. It allows users with Macintoshes at home and with Apple Remote Access versions 1.x or 2.0 to directly access their Macintosh resources at the Lab. When dialing in, just respond to the `t-s-modem` prompt with `ARAP`. There is no support for the so-called "magic character" when making the connection. One resource for modem scripts for ARAP is John Urish of the Computing Division's Personal Computing Support Group. John may be contacted at x3017.

Direct support for ARAP V2.0 will be made available as soon as it completes beta testing.

More Dial-in Lines Added

Usage patterns of the direct dial lines (840-8134) vary somewhat with the season of the year. Winter months show the heaviest use with regular 85-90% peak use during evening hours. 24 additional phone lines will be added in September to bring the total available to 72. No additional call-back lines are planned at this time.

Andy Lego, x2961, andylego@fnal.gov

Al Thomas, x3064, thomas@fnal.gov

VAX

FNALV Futures: A Mixed VAX-AXP Architecture Cluster

The VMS Systems Support (VSS) group plans to upgrade the Central Facility VAX Cluster FNALV early this fall. Other on-site clusters, notably FNALD and FNALD0, will be upgraded shortly thereafter. The upgrades will include the addition of an AXP, a new computer platform running VMS. This will necessitate a reorganization of the applications software supported on FNALV (and elsewhere) to accommodate a mixed-architecture (VAX and AXP) cluster. It will also require an upgrade to OpenVMS V6.1, the first stable release of VMS which is fully compatible between the two architectures. Several other system software products (most notably, the compilers) will need upgrading during this time frame as well.

In 1992 Digital Equipment Corporation formally introduced a new platform of computer which is able to run the VMS operating system. This new platform, known alternately as the Alpha or AXP system, provides much higher compute power and I/O throughput than the VAX platform. To the end user, the AXP is essentially identical to the VAX, with a common DCL interface and system service support. To the programmer or application support person, however, there are underlying differences between the two architectures which need to be considered when developing, porting, or maintaining VMS applications.

The main issue in supporting a mixed-architecture cluster is that object modules, libraries, and executable images – in short, all binary program files (not data files) – are incompatible between the two platforms. To minimize confusion and inconsistency in behavior between nodes of the cluster, a means of distinguishing between binary files produced on the two platforms must be established.

Many ideas have been proposed, and several methods are currently in use at other sites. None have, as yet, become "industry standards". After studying the issue, VSS has decided to follow the following scheme:

- Source code will be maintained, as much as possible, in common directories such as `LIB:[LIB.product.CURRENT]` (for products in `LIB`) or `product$ROOT:[SOURCE]` (for products in `SITE_PRODUCTS` format).
- Architecture-specific files such as object libraries, executable images, and include files, will be maintained under the appropriate subdirectory, `[.VAX]` or `[.AXP]`. In some cases, this will need to be further extended to, e.g., `[.VAX_OLB]` and `[.AXP_OLB]`, `[.VAX_EXE]` and `[.AXP_EXE]`, etc. The important thing to note is that both architectures will be clearly spelled out in the naming convention of platform-specific files.
- `SETUP.COM` files will determine the current architecture and define logical names and/or symbols to point to the appropriate files.

For example

```
$IS_AXP = F$GETSYI("NODE_HWTYPE").EQS. "ALPH"
$IF(IS_AXP)
$THEN
$DEFINE product_EXE -
    product$ROOT:[AXP_EXE]product.EXE
$    DEFINE product_OLD -
    product$ROOT:[AXP_OLB]product.OLB
$ELSE
$    DEFINE product_EXE -
     product$ROOT:[VAX_EXE]product.EXE
$    DEFINE product_OLB -
    product$ROOT:[VAX_OLB]product.OLB
$ENDIF
```

By using this approach we may continue to provide a consistent interface to supported applications software, whether on a VAXcluster, an AXP cluster, a stand-alone system of either variety, or a mixed-architecture cluster. This product support structure will not need revision when AXPs are added to an existing VAXcluster, nor will it need modification when all VAXen are retired from a mixed-architecture cluster leaving behind only AXPs. We recommend that other code maintainers on site include this scheme in their considerations when they address the AXP issue.

Over the next few months the Physics Analysis Tools (PAT) group, Mike Stolz, and I (among others) will be porting and/or upgrading the `LIB` and `SITE_PRODUCTS` suite of supported products. Not all software will be ported to AXP, especially in the initial stages. Several software packages which are no longer in use will be retired at this point. Details will always be posted in **INFO**.

Lauri Loebel Carpenter, x2214, lauri@fnal.gov

Schedule for the AXP (Alpha) Upgrades for the FNALD and FNALV Clusters

"Modernizing the FNALV Cluster" in the previous Computing News introduced the plan to upgrade the FNALV cluster with the addition of a DEC 7630 AXP system. This system is now on-site and should be installed by the time of the publication of this issue of the Computing News.

Initially the system will be not be available for general users. Computing Division personnel will be actively working on setting up the system and porting various VMS products to the AXP architecture. During this time, the AXP system will be part of the FNALV cluster but will not respond to access via the FNALV cluster alias node name.

It is expected that some 6-to-8 weeks after the system is installed, we have a suite of VMS applications available on the AXP that will allow a significant percentage of the FNALV user community to effectively use the AXP system. It is expected that we will be able to open the AXP up to the user community in late September or early October of 1994. At this time, the AXP system would be accessible via the FNALV cluster name.

We do not expect to have ported all the planned VMS products by this time so some users will find that they will need to continue to use the VAX systems of the FNALV cluster until the products they are dependent upon become available on the AXP system. Users should watch **INFO** for the announcements of VMS products available on the AXP architecture.

We hope to complete the planned set of ports by Thanksgiving (end of November 1994) and have most of the FNALV users able to effectively utilize the AXP system. There will be a few products (mostly commercial products from third-party vendors) which will remain available only on the VAXes in FNALV and will not be usable on the AXP system.

At the same time the FNALV AXP system is being readied for users, a pair of Digital 2100 (Sable) multiprocessor AXP systems will be installed in CDF's FNALD cluster. These systems will also depend upon the VMS product porting work being done on the FNALV AXP system. The schedule for these systems to be available for Early Bird CDF users and then the entire CDF community should roughly match the above schedule laid out for the FNALV AXP system.

Soon after the AXP system becomes available for FNALV users, the FNALF (VAX 8830) system will be removed from the FNALV cluster. This system is being returned to Digital as part of the AXP upgrade program. At some further point, as more FNALV users have switched to the AXP system, the FNALO (VAX 6550) system will be removed from the FNALV cluster and transferred to the FNALD cluster to augment its capacity. In addition, more of the FNALNx systems (the MicroVAX 3100s) will be removed from the FNALV cluster and retired.

Frank Nagy, x4935, nagy@fnal.gov

Potentially Obsolete VAX Products

Over the last ten years, over 200 different VAX/VMS software packages have been made available by the Computing Division to Fermilab users on the FNALV and FNALD clusters and to some extent FNALD0. These products, in `LIB` or `SITE_PRODUCTS` format and accessible via the **SETUP** command, have made life much easier for our VAX users. However, many of these products are now rarely if ever used as currently installed on these systems and are not worthwhile supporting in this way as we move towards Unix and Alpha systems.

A major effort is underway to identify these little-used products. Some serve no purpose since the hardware they supported no longer exists at Fermilab. Others are more suitable on UNIX systems or are now supported in other ways. For example, the individual CERN products are supported in the **CERN** product. Yet others have been replaced with better software tools, or just aren't adequate for our needs any more.

We do plan to retain and support all VMS products that still serve a useful purpose, even if they are not heavily-used, but little-used products may never become available on the DEC Alpha systems

Please take a look at the following list of `LIB` and `SITE_PRODUCTS` software. These products have all been flagged as low or no-use products. If you need and use any of them, please send an e-mail message to `stolz@fnal.gov`, letting me know, briefly, why and how often you use the product.

* HALO	* HOST	~ MONTERM
* PATCHY	* SIX12	* MULANA
* TERMTABLE	* TRANSPORT	~ PREPRINTS
* TURTLE	* VFORTCD	* SAGE
* VM	* VMTOOLS	~ SDCNEWS
* ANSIFOR	~ BITNET	~ SPLIT
~ BULLETIN/TELL	+ CYBER	+ WYLVAX
* DECUS	~ EGS	~ MORTRAN
* EXPORT	~ FLUIDS	* NOTICE
~ FORTTABS	~ GKSPRT	~ PUBHELP
~ GUIDE7	* HANDYPAK	* SCRFT
* HPLOT4	* IMPORT	* SLATEC
* JETNET	* KIOWA	~ TEXTOOLS
* ZIPMAP		

* `LIB` product not used?
~ usage being logged - minimal use of `LIB` product?
\+ being watched closely

Mike Stolz, x8060, stolz@fnal.gov

DART V2

A major milestone of **DART**, the data acquisition system (DA) for the next fixed target run, has recently been achieved and released as **DART** V2. V2 includes the first release of the **DART** tape logging and run control software for Unix, and level 3 event building software. The first release of **DART** was discussed in the August - October 1993 newsletter.

E811 has used the first release of **DART** to take data since the beginning of the year. Since then, as a result of the collaborative effort between the Online Systems Department and E781, E835, E831, E815, and KTeV, DART test systems have been established for each experiment with data flowing from a front end crate in CAMAC or FASTBUS through to the backend "Level 3" and analysis workstation using **DART** hardware and software components. Single stream test stands, using **DART** V2 software, have also been constructed in the OLS 9th floor crossover area.

Available as part of **DART** V2 are several new products as well as upgrades to existing hardware and software. Due to limited space in this month's newsletter, mostly new software will be described. A description of upgraded software will be included in the next newsletter. Upgraded software includes buffer manager software, **dfm**, and coupled event distribution software, **dfm_hoist**; real-time operating system software for embedded processors, **vxworks**, including support tools and board support packages, **vx_tools**, **vx_dart**, **vx_fscc**; and FASTBUS standard routines for the FASTBUS Smart Crate Controller **fb_std** and **fb_fscc**.

DART V2 is available for IRIX V5.1 (IRIX V5.2 is in test) and SunOS V4.1. Full documentation can be found in the **DART** home page on WWW. (The URL is `http://fndaub.fnal.gov:8000/`). A list of the acronyms used in this article is available on this home page. All software can be distributed via the `kits` database on `fnsg01.fnal.gov` using **upd** or from `ftp.fnal.gov`.

Ruth Pordes, x3921, ruth@fnal.gov

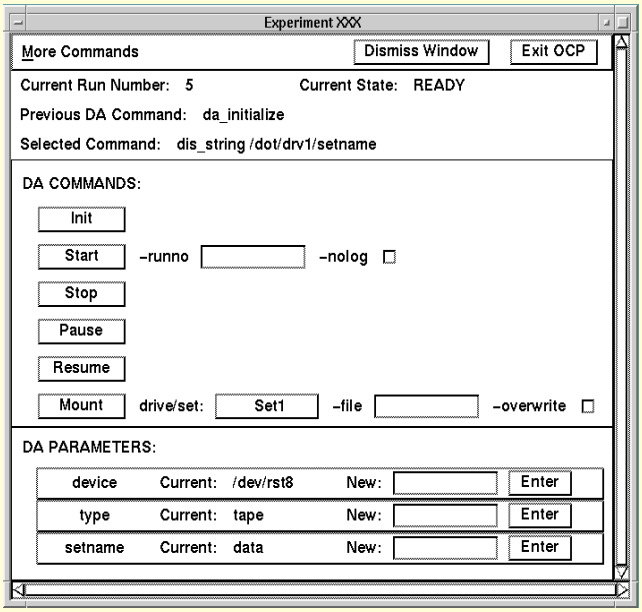
Gene Oleynik, x2430, oleynik@fnal.gov

C D

New Host Software

Run Control

Run control is a suite of DA host software used to control applications on various computers, including level 3 event building, filtering, and logging, as well as applications running on front-end readout controllers. DART Run Control provides an Operator Control Program (**ocp**) through which these distributed applications can be controlled. The **ocp** uses the **tcl** command line interpreter through which users can invoke existing run control commands, tailor them, or add their own commands by editing a startup script. A graphical interface is in the making (see graphic below).



Through run control, applications like tape loggers define and join "groups", for instance, a "logger" group. The DA operator can then send commands to the "logger" group with the **ocp**, and all applications that have joined the "logger" group will receive the specified command. The usual set of commands are provided (`start`, `stop`, ...), which send the commands to appropriate groups in the appropriate order. This group concept allows a "canned" DA to be defined which is easily adapted to an experiments specific needs.

A client library provides an interface for applications to join and leave groups and to send or receive run control commands and replies. Both C and **tcl** calls are provided so either C or **tcl** language applications can participate in run control.

Another important aspect of run control is providing configuration parameters to the applications that make up the DA. **DART** provides a distributed system which allows users to configure their DA by specifying run control parameters and values in a configuration "database". The database can be loaded and unloaded from ASCII files. Both C and **tcl** calls are provided to store and fetch parameters. This configuration mechanism allows experimenters to tailor **DART** for their own use with minimal or no changes to their C/Fortran code.

The software products that make up run control are: **dbs** (Dart Bootstrapping Services) for starting DA applications from a single node, **drc** (Dart Run Control client library) for sending and receiving run control commands, **dms** (Dart Multicasting Services) for handling group multicasting over the network, **dis** (Dart Information Services) for storing and retrieving DA configuration parameters, and **ocp** for operators to control the data acquisition system. **DART** V2 requires the following

versions: **dis** v1_3, **dbs** v1_1, **drc** v1_0, **dms** v1_1, and **ocp** V1_0.

Gene Oleynik, x2430, oleynik@fnal.gov

Laura Appleton, x8107, appleton@fnal.gov

Carmenita Moore, x2288, moore@fnal.gov

Lourdu Udumula, x2215, udumula@fnal.gov

DART Integration – dart v2_0

The **DART** DA consists of not only the applications such as the logger, but host servers which support the distributed nature of the system. These servers and various other **DART** V2 components have been integrated so that the user can start them up in a simple manner. The software product that supports this is named **dart**.

The **dart** product provides a script file that defines two **tcl** procedures which are invoked from the dart bootstrap program (**dbs**, discussed in a previous newsletter) to start the DA up:

- **boot_host** – Boot the host servers on the host node, and start the operator control program (**ocp**) from which the DA can be controlled.
- **boot_target** – Boot the application code, including the tape logger (**dot**), on another or the same node. **boot_target** can be invoked more than once to boot applications on any number of nodes.

These tools are provided to aid in starting up V2 systems, and will probably change in form for future releases; one of the goals for **DART** V3 is to streamline the startup of DA applications.

Eric Van Drunen, x3635, drunen@fnal.gov

New Level 3 Support Software

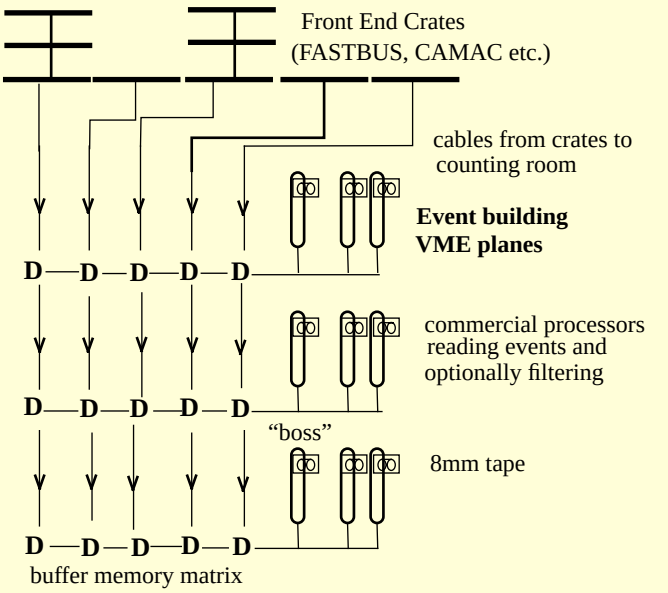
ddd v2_3

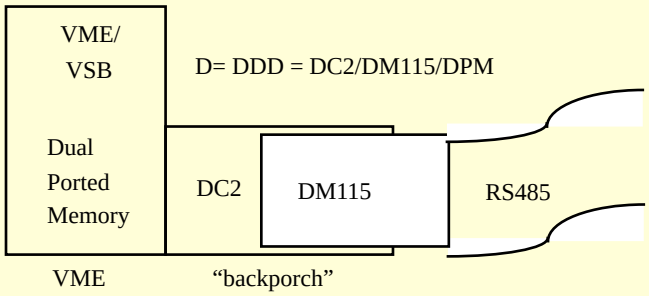
The DDD is an intelligent dual ported memory for level 3 buffering, and is a key component of the DART event building architecture. A DDD is constructed from 3 integrated modules: DC2, DM115, DPM (see graphic below). The DC2 is an embedded processor that controls and manages the data pushed from front-end readout modules (via a DM115) into the VME/VSB dual ported memories (DPM). DDDs are represented as Ds in the DART architecture graphic.

DDDs are fed sub-event data over DART standard RS-485 cables from FASTBUS with the FASTBUS Smart Crate Controller, CAMAC via DYC+ or CAMAC Smart Crate Controller, or from Fera or PCOS via DYC+ readout controllers.

DART V2 contains two software products for the usage and tests of the DDD Dart component: **ddd** and the **dddmbx**.

The **ddd** product contains the firmware that runs in the DC2 modules to manage the dual ported memories. It receives sub-event data from the experiment front end sources, stores them in the dual ported memories, and manages a table that permits





random access to the event data over VME. The DC2 has a Motorola 68340 microprocessor with a DMA engine, and communicates with a VME master via a mailbox in the dual ported memory.

The **dddmbx** product is a utility that runs on an MVME167, a single-board VME computer running VxWorks, to let the user enter commands in the DC2 mailbox, monitor DC2 activity, and read events deposited in the DPM by the DC2.

The function of the **dddmbx** software is to test and diagnose the DDD **DART** component. The **ddd** firmware will read out events during the run and commissioning phases of experiments.

Oscar Trevizo, x3257, trevizo@fnal.gov

Data Readout into UNIX – gateway v1_0

Once sub-event data is collected in the dual ported memories in the VME "event building" crate, they need to be analyzed for event selection in the experiment specific Level 3 software. For some experiments, the Level 3 software needs to run in powerful processors with support for high level languages and debugging.

In DART, we are using the multi-processor Silicon Graphics Challenge L computers, which have high bandwidth CPU, Memory and I/O capabilities. The data is transferred from the VME crate directly into Challenge memory via PT-940VME or

BIT3 VME-to-VME links which are capable of transfer rates of up to 40 MByte/sec. Once in memory, the sub-events are made available to one of the Challenge processors. The single program which reads and clears the VME memories, and distributes the event data to the Level 3 processors is called **gateway**.

gateway uses low level driver software to access the event building VME crate. Drivers are provided to control the PT-940VME and the BIT3 VME-to-VME links. The PT-940VME has high-rate VME64 capability but relatively high per transfer overhead, while the BIT3 does not have VME64, but has lower overhead. In order to be able to switch between these different links, DART contains a software layer called DART Data Transport (**ddt**), which is a library of common functions to transfer data between the Challenge and VME crates independent of the link type.

The PT-940 VME device driver was developed as part of the SDSS project, and the BIT3 driver was written specifically for the DART project. Both drivers have been run extensively on the Challenge with IRIX 5.1.1.1. The PT-940 VME driver is also supported on the Crimson and 4D/35 running IRIX 4.0.5, and the BIT3 is supported on the Crimson.

Jonathan Streets, x3629, streets@fnal.gov

Jurgen Engelfried, x5151, jurgen@fnal.gov

Tape Logging – dot v1_2

Events accepted by the level 3 filters need to be logged to tape. **dot** (data on tape) is a logger for recording to tape or disk files. **dot** is integrated with the DART buffer manager (**dfm**). It currently supports Exabyte 8500s, and can log to any number of drives across multiple SCSI buses. Drives can be organized into drive sets, which permit logging to drives in parallel, thus exceeding the rate capacity of a single drive. Applications make **dfm** logging requests to **dot** by specifying a drive set. By this means, different types of events can be routed to different sets.

Various configuration information about logging is stored in the **dis** configuration database. The logger itself is controlled by run control, processing commands to initialize, mount, start, and stop.

DART specifies two standard formats for writing to tape and disk: record and event level. Both support blocking events into records, which is required to keep the 8mm drives streaming at full speed. The only requirement placed on event formatting is that the first 32 bit word in an event contain the inclusive byte count. These standards are optional, but support is provided for their use.

dot is available for IRIX and SunOS systems.

Margaret Votava, x2625, votava@fnal.gov

Tape Reading – daft v1_0

dot has a companion product to read back "**DART** formatted" tapes or disk files, called **daft** (data from tape), for Unix or VMS systems. **daft** supports access to multiple devices from a single program, so once the event is read, **daft** can also be used to write it back out to a different tape drive or disk.

daft is organized in two layers. The top layer deals only with events, and unblocks DART formatted events. The bottom layer deals only with records, and most all device/operating system dependence is contained in this layer. **daft** is meant to be easily transportable to other Unix systems. **daft** has both C and Fortran calling interfaces.

Francesco Prelz, x2183, prelz@fnal.gov

C D

New and Upgraded Hardware

The FASTBUS Smart Crate Controller, or FSCC, module hardware has been upgraded to more effectively match the **DART** architecture. Known FSCC-PC4 bugs were fixed, and enhancements including larger RAM and ROM memory spaces, new microcode instructions, zero word event suppression, and buffering for both the front-panel Trigger and Trigger-ID bits have been added. Modifications to the FSCC Output Port Controller were made to comply with the **DART** protocol. The FSCC-PC4A is available from Bira Systems.

To accommodate the required cable length needed at E781, an "RS-485 extension cord" has been developed in the form of a fiber optic transmitter mother-module hosting a commercial fiber optic transmitter daughter submodule, and the matching receiver mother/daughter pair. These form a transmitter/receiver set. The modules are currently being commissioned in the E781 DA system.

DART V2 additionally supports data readout through the DYC+ - a CAMAC-FERA readout controller under development by the Physics Section.

A single stream **DART** test stand on WH9X is being maintained to provide an available and understood test stand for modules that may not behave as expected in experiment systems.

Bob Forster, x2535, forster@fnal.gov

C D

Data Flow Management – dfm v2_7_2

The DART data flow manager, **dfm**, the buffer manager for DART previously only available under the VxWorks operating system, has been ported to the SunOS and IRIX operating systems. Shared memory segments are used for control structures and either a shared memory or kernel memory segment for data buffer areas. A select function has been added, which allows a provider process to wait on multiple **dfm** message queues. A wait forever option has been implemented which allows an application to block on a message queue with no associated time-out.

David Berg, x3021, berg@fnal.gov

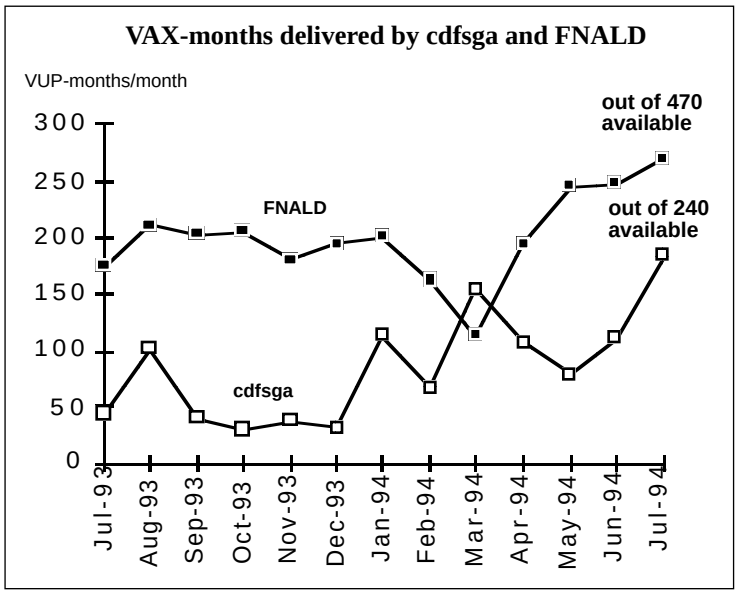
Dennis Black, x4525, dblack@fnal.gov

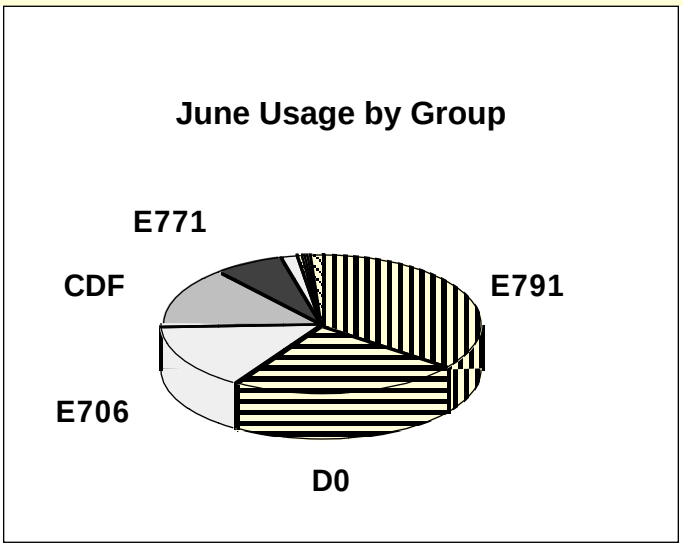
Computer Usage

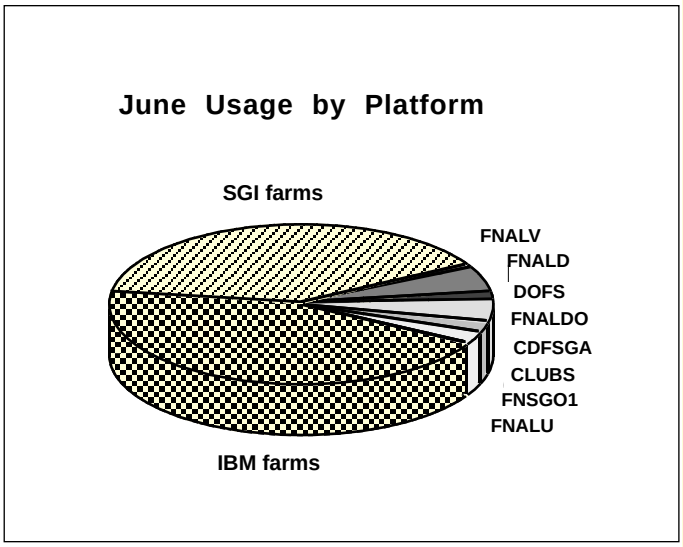
The graph at the right plots CDF usage of FNALD and cdfsga in VUP-months/month.

The charts and table below summarize the computer usage of the centrally-supported systems in the month of June. The pie charts break down the usage by group/experiment and by platform. The table displays the usage of the largest groups by platform in VUP-months.

Judith Nicholls, x3989, nicholls@fnal.gov



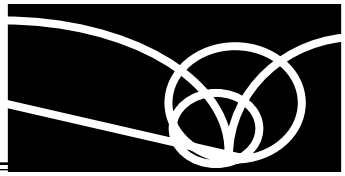




Accounting data for June 1994

Group	FNALV	FNALD	D0FS	FNALD0	CDFSGA	CLUBS	FNALU	FNSG01	FRM-IBM	FRM-SGI	Total
E791	0	0	0	0	0	0	0	0	1443	270	1713
D0	0	0	70	238	0	2	0	0	14	785	1110
E706	1	0	0	0	0	0	0	0	0	762	763
CDF	1	238	0	0	85	0	0	0	309	18	651
E771	0	0	0	0	0	0	0	0	315	8	324
Acc Theory	0	0	0	0	0	90	2	0	0	0	92
E687	1	0	0	0	0	0	0	0	0	13	15
E761	0	0	0	0	0	14	0	0	0	0	14
E665	0	0	0	0	0	9	0	0	0	0	10
E672	10	0	0	0	0	0	0	0	0	0	10
E783	5	0	0	0	0	0	0	0	0	0	5
Other	16	9	0	0	8	7	2	4	4	7	57
Total	35	247	70	238	93	122	5	4	2085	1863	4763

Change of Address



If you have changed your address, or are not on *Computing News* mailing list and would like to placed on it, or would like to be removed from the list, please complete this form and mail it to the Editing Staff.

Action to be Taken (check one):

Deletion ❑ Addition ❑ Correction ❑

Return completed form to:

Pat Haring
Fermilab - MS 120
P.O. Box 500
Batavia, IL 60510

Address Information (Include Zip Code, please)

_____

_____

_____

_____